

## **ACCURACY ASSESSMENT OF RESPONSE SURFACE APPROXIMATIONS FOR SUPERSONIC TURBINE DESIGN**

Nilay PAPILA, Melih PAPILA, Wei SHYY, Raphael T. HAFTKA and Norman FITZ-COY

Department of Aerospace Engineering, Mechanics & Engineering Science  
University of Florida, Gainesville, FL

### **1. INTRODUCTION**

There is a growing trend to employ CFD tools to supply the necessary information for design optimization of fluid dynamics components/systems. Such results are prone to uncertainties due to reasons including discretization errors, incomplete convergence of computational procedures, and errors associated with physical models such as turbulence closures. Based on this type of information, gradient-based optimization algorithms often suffer from the noisy calculations, which can seriously compromise the outcome. Similar problems arise from the experimental measurements.

Global optimization techniques, such as those based on the response surface (RS) concept are becoming popular in part because they can overcome some of these barriers. However, there are also fundamental issues related to such global optimization technique such as RS. For example, in high dimensional design spaces, typically only a small number of function evaluations are available due to computational and experimental costs. On the other hand, complex features of the design variables do not allow one to model the global characteristics of the design space with simple quadratic polynomials. Consequently a main challenge is to reduce the size of the region where we fit the RS, or make it more accurate in the regions where the optimum is likely to reside. Response Surface techniques using either polynomials or and Neural Network (NN) methods offer designers alternatives to conduct design optimization. The RS technique employs statistical and numerical techniques to establish the relationship between design variables and objective/constraint functions, typically using polynomials. The NN technique employs many simple linear and non-linear elements operating in parallel and connected in patterns to represent

such relationship between design variables and objective/constraint functions. The polynomial and NN techniques can be used either independently or in combination. Depending on the characteristics of the design variables, polynomials and NN can exhibit different accuracies in different regions of design space. Hence, a main interest of the present effort is to identify ways to combine polynomial and NN techniques to enhance the performance of the overall RS model.

In this study, we aim at addressing issues related to the following questions: (1) How to identify outliers associated with a given RS representation and improve the RS model via appropriate treatments? (2) How to focus on selected design data so that RS can give better performance in regions critical to design optimization? (3) How to combine NN and polynomial techniques for improving the accuracy of the RS model?

## **2. MAIN APPROACH AND SCOPE**

The physical example chosen in the present study is the supersonic turbine envisioned for the next generation reusable launch vehicle (RLV). There are growing interests to consider this technology for space transport. Based on our previous work [1-3], a two-stage configuration has been optimized at the preliminary design level. The focus here is to optimize the shape of the stator (vane) and runner (blade) in each stage. Navier-Stokes-based CFD solutions are used as the sole input data. For the first stage vane, there are 7 design variables, while for the first and second stage blade and second stage vane, there are 11 design variables. In all cases, the goal is to maximize the stage total-to-total efficiency ( $\eta$ ).

### **2.1. Outlier and Bias Error Analysis**

We intend to identify the data points that are "statistically" out of the range for the response surface (RS) model under consideration and characterize them as *outliers*. *Outliers* are defined as infrequent observations that do not appear to follow the characteristic distribution of the rest of the data and they may have a strong influence on the least squares estimate. Statistical analysis can be utilized to detect such flaws.

- ***Outlier Analysis based on Iteratively Re-weighted Least Square*** (IRLS) procedure will be adopted for detection of the outliers [4, 5]. We hope that detecting outliers will help us to offer more insight into following problems

- A better understanding of the scatter of the data generated directly by CFD.
- The effect of the outliers on the calculation of statistics and degree of fidelity of the response surface model. The number of outliers can indicate the degree of fidelity of the RS.
- How to interpret and handle such design points for the given application problem. Excluding all outliers might not be the best solution especially if the nature of the outlier design is not clear.

Statistical tools and associated assumptions may also introduce additional uncertainty. Therefore, we are going to use an alternative approach called as *Mean Square Error-Based Approach* together with an *Outlier Analysis* while searching for the ways of defining uncertainties associated with the generated response surface model.

- A ***Mean Square Error-Based Approach*** addressing the approximation errors due to model inadequacy will be applied [6]. The approach seeks to determine locations in the design space where the accuracy of the approximation appears poor. This approach can help to assess the certainty of predicted optimal designs.

## **2.2. Selective Emphasis of Critical Input Data**

Since we are most interested in identifying highest efficiency points, using the outlier analysis and mean square error approach, we can place higher emphasis on data belonging to such a region [4, 7] to improve the model performance in critical areas and/or identify needs for further

input data. For example, we can assign higher weightings for data with higher efficiency values when applying the IRLS approach. Also, the level of scatter for training and testing points close to design goal can be calculated to illustrate the expected uncertainty of the RS prediction.

### **2.3. NN-Enhanced RS Model**

In our previous research, it is demonstrated that to use the information obtained by using outlier analysis and mean square error approach to select the design points to be generated additionally using neural networks. This approach is often applied to supply additional information for the polynomial response surface by using *Neural Network* (NN) trained by the original CFD data. This can be used to improve the accuracy of the RS, and to allow the optimization task to be conducted with smaller number of CFD runs. Ultimately, we want to see how to use effectively neural networks and different level of response surface to maximize the performance of the optimization tool. However, there are few critical issues that need to be focused on when creating such *NN-Enhanced design space*.

- The distribution of the data to be added using NN's, for example, should be selected systematically. It can either be chosen in such a way that it fills-out the "holes" of unrealistic or difficult cases for which CFD tools may not be suitable, or it can follow one of the DOE techniques that is going to enrich the original design space in a more systematic way.
- The ratio of the number of original data (CFD) and enhanced data (NN) can have an effect on response surface efficiency. For example, if the number of enhanced data generated by NN is much larger than the original CFD data, this might overwhelm the characteristics of the problem.

## **3. PRELIMINARY RESULTS**

We have considered the first vane shape design optimization of a supersonic turbine as an application problem. For this case, there are 7 design parameters and the objective function is the

stage total-to-total efficiency ( $\eta$ ). For this case, CFD information is available at only 245-design points that are reduced from face centered composite designs in  $(-1, +1)$  for all design variables. Among these 245-data, 219 of them are used for fitting and the remaining 26-data is used for testing the approximation accuracy that we constructed. The range for the test set is  $-0.5$  to  $0.25$ .

We have studied 3 quadratic approximation models: (1) RS without outliers treatment, (2) Standard IRLS, and (3) IRLS customized by higher weight assignment to data in high-efficiency design regions (We define a high-efficiency design if  $\eta \geq 0.75$ ). The main difference between the last two models is the weight distribution used for IRLS. In standard IRLS procedure, the weight distribution given below is used and 2<sup>nd</sup> model assigns the weights according to this formula. However, for the customized model, weights are forced to be not lower than 0.8 for designs of  $\eta \geq 0.75$  region.

$$w = \begin{cases} \left[ 1 - \left( \frac{\epsilon/\sigma_a}{B} \right)^2 \right]^2 & \text{if } |\epsilon/\sigma_a| \leq B \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The statistical summaries of these models are shown in Table 1. Figure 1 illustrates the performance of the original RS and IRLS models along with the outliers, based on CFD-data. Together with Table 1, it shows that by treating the outliers, better models can be constructed. Since we are ultimately interested in determining an optimal design, it is instructive to check the range of scatter as marked on Figure 1 (b) and (c) associated with the original CFD data. Table 2 compares the number of outliers contained in either approach. Standard IRLS detects 17 outliers with 7 existing in the higher efficiency region. Customized IRLS, however, finds 15 outliers with all existing in the lower efficiency region as expected. Figure 2 shows results from mean squared error criterion based approach for the quadratic RS approximation. The approach presents a point-wise measure (eigenvalues) characterizing possible bias error assuming a cubic model as the true function. Positive correlation between the eigenvalues and the magnitude of bias error is expected

in case the fitting model is inadequate. We use absolute error between the CFD data and the quadratic RS predictions for the evaluation. We also checked the correlations between the efficiency and the eigenvalues/error in order to investigate the modeling error distribution in high and low efficiency design regions and reported in Table 3. Negative correlation between the efficiency and the errors is an indication that the quadratic RS is predicting better for high-efficiency region although all data points have same weight=1.

#### 4. REFERENCES:

1. Papila, N., Shyy, W., Griffin, L., Huber, F. and Tran, K., "Preliminary Design Optimization for a Supersonic Turbine for Rocket Propulsion," *AIAA/SAE/ASME/ASEE 35th Joint Propulsion Conference*, Paper No. 2000-3242, July 16-19 (2000).
2. Papila, N., Shyy, W., Griffin, L., and Dorney, D.J., "Shape Optimization of Supersonic Turbines Using Response Surface and Neural Network Methods," *AIAA 39th Aerospace Sciences Meeting & Exhibit*, Paper No. 2001-1065 (2001).
3. Shyy, W., Papila, N., Vaidyanathan, R. and Tucker, P.K., "Global Design Optimization for Aerodynamics and Rocket Propulsion Components," *Progress in Aerospace Sciences*, Vol. 37, (2001) pp.59-118.
4. Papila, M. and Haftka, R. T., "Response Surface Approximations: Noise, Error Repair and Modeling Errors," *AIAA Journal*, 38(12), pp. 2336 – 2343, 2000.
5. Kim, H., Papila, M., Mason W. H., Haftka, R. T., Watson, L. T., and Grossman, B., "Detection And Correction Of Poorly Converged Optimizations By Iteratively Reweighted Least Squares," *41<sup>st</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference*, Paper No. 2000-1525, April 3-6 (2000).
6. Papila, M. and Haftka, R. T., "Uncertainty and Response Surface Approximations", *42<sup>nd</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Paper No. 2001-1680, April 16-19 (2001).
7. Balabanov, V. O., Giunta, A. A., Golovidov, O., Grossman, B., Mason, W.H., Watson, L.T., and Haftka, R.T., "Reasonable Design Space Approach to Response Surface Approximation," *Journal of Aircraft* 36(1), 1999, pp. 308-315.

**Table 1.** Statistical Summaries of different quadratic models constructed for the **first vane**

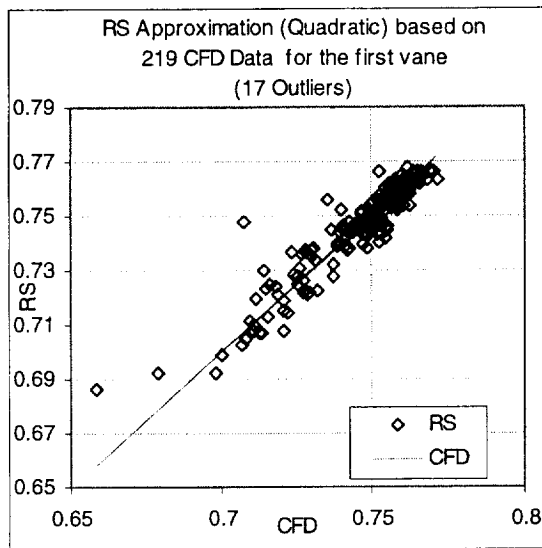
	<b>RS for CFD Data</b>	<b>Standard IRLS</b>	<b>Customized IRLS</b>
RSquare	0.879	0.951	0.937
RSquare Adj	0.856	0.941	0.924
Root Mean Square (rms) Error	0.007	0.004	0.004
%rms-Error	0.874%	0.490%	0.536%
Mean of Response	0.747	0.749	0.750
Observations (or Sum Wgts)	219	202	204
Testing rms-Error	0.003	0.003	0.001
% Testing rms-Error	0.437%	0.418%	0.174%

**Table 2.** Outliers Summary for different quadratic models for the **first vane**

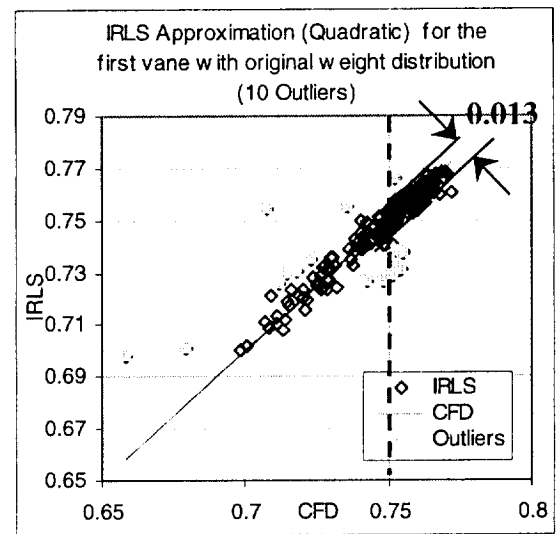
	Number of Outliers in CFD Data	Number of Outliers in $\eta < 0.75$	Total Number of Outliers	Total Number of Data
<b>RS for CFD Data</b>	17	10	17	219
<b>Standard IRLS based on CFD data</b>	-	-	-	202
<b>Customized IRLS based on CFD data</b>	15	15	15	204

**Table 3.** Coefficient of correlation summary for the **first vane**

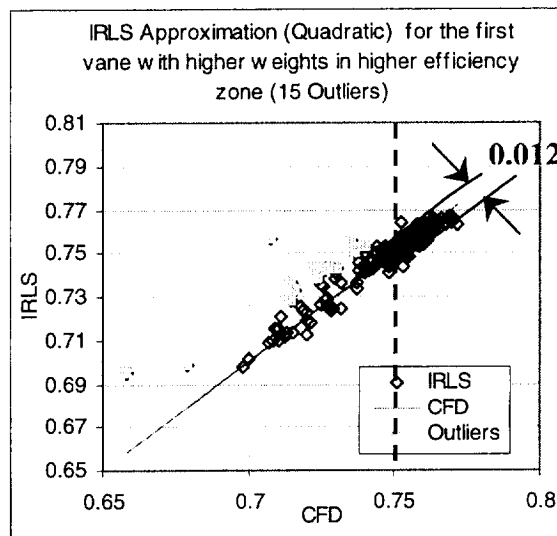
	<b>Eigenvalues</b>	<b> Error </b>	<b>%  Error </b>	<b>Efficiency</b>
<b>Eigenvalues</b>	1.000	0.227	0.231	-0.465
<b> Error </b>		1.000	0.999	-0.451
<b>%  Error </b>			1.000	-0.476
<b>Efficiency</b>				1.000



(a) RS model



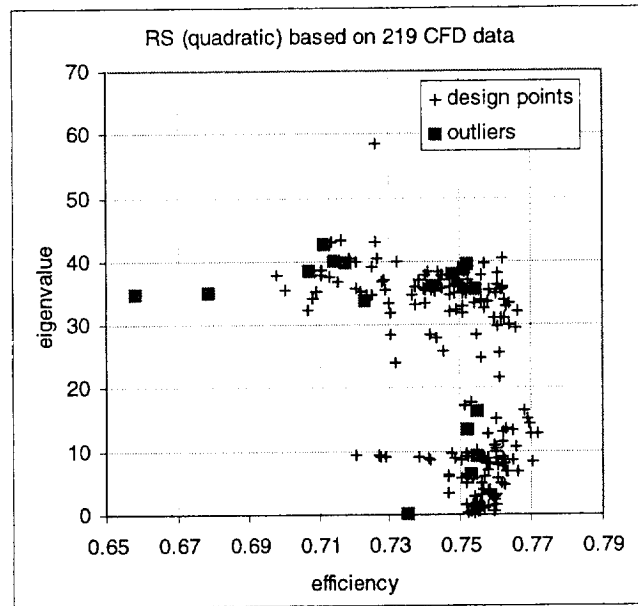
(b) IRLS with original weight distribution



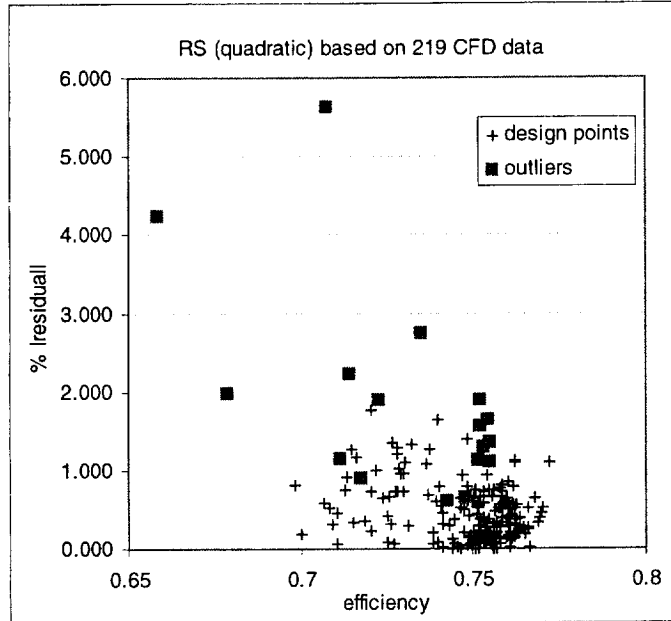
(c) IRLS with modified weight distribution

**Figure 1.** Leverage plots for Efficiency with 3 different models based on CFD-Data only for the first vane

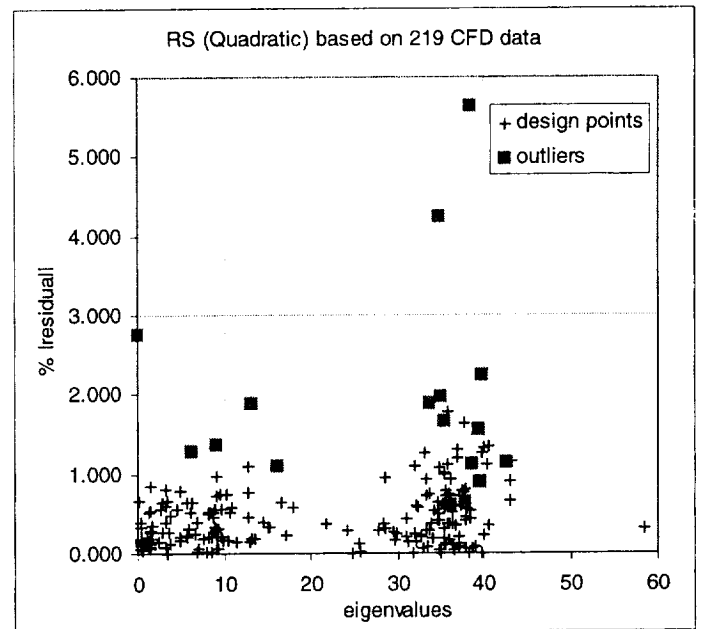




(a) Efficiency vs. Eigenvalues

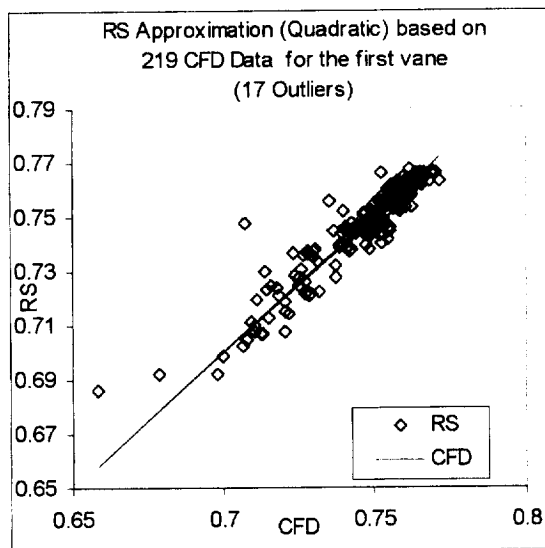


(b) Efficiency vs. % |Error|

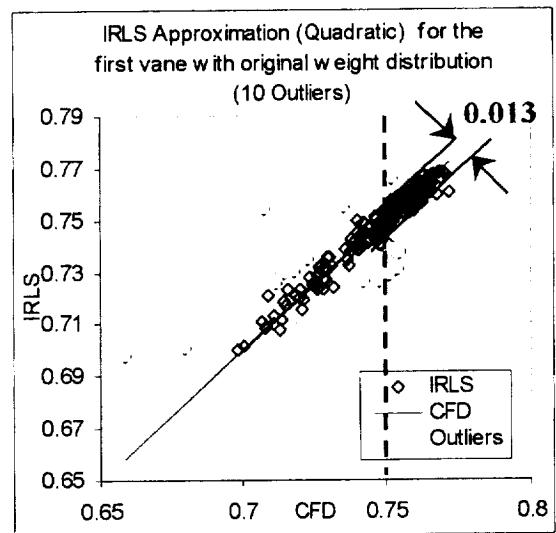


(c) Eigenvalues vs. % |Error|

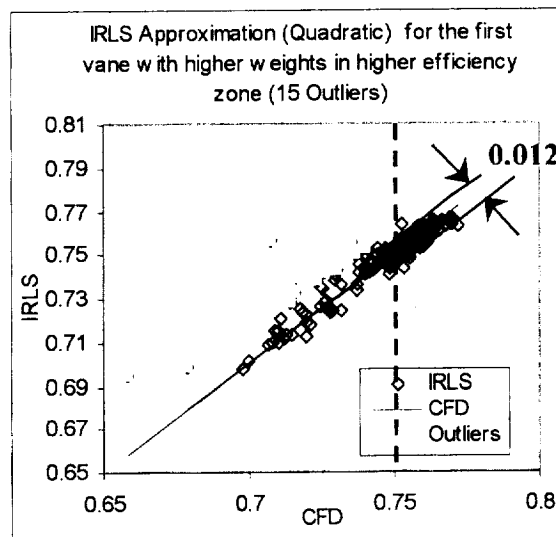
**Figure 2.** Quadratic RS bias error analysis against cubic RS by **Mean Squared Error** criterion based on CFD-data for the **first vane**



(a) RS model

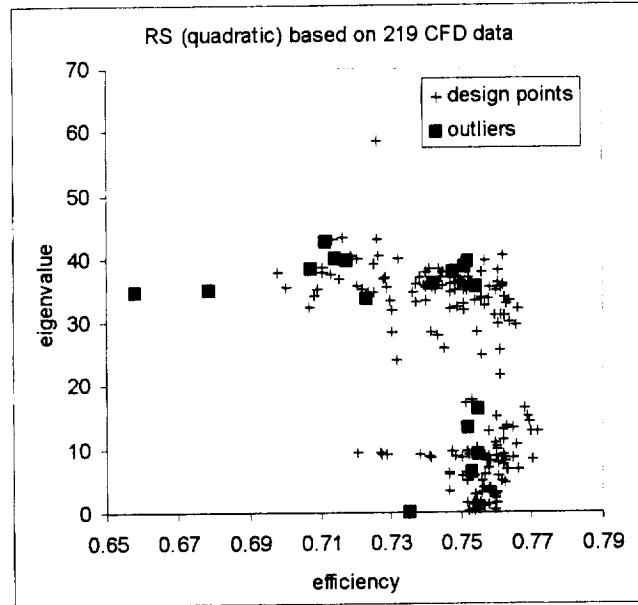


(b) IRLS with original weight distribution

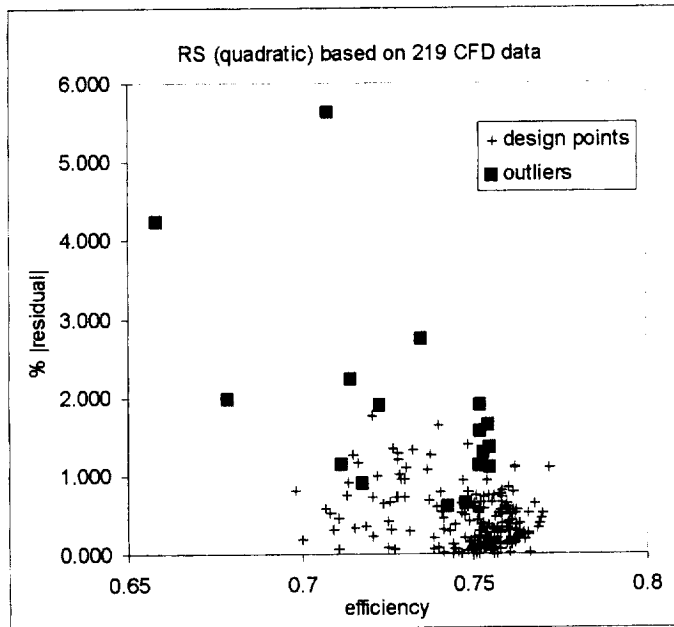


(c) IRLS with modified weight distribution

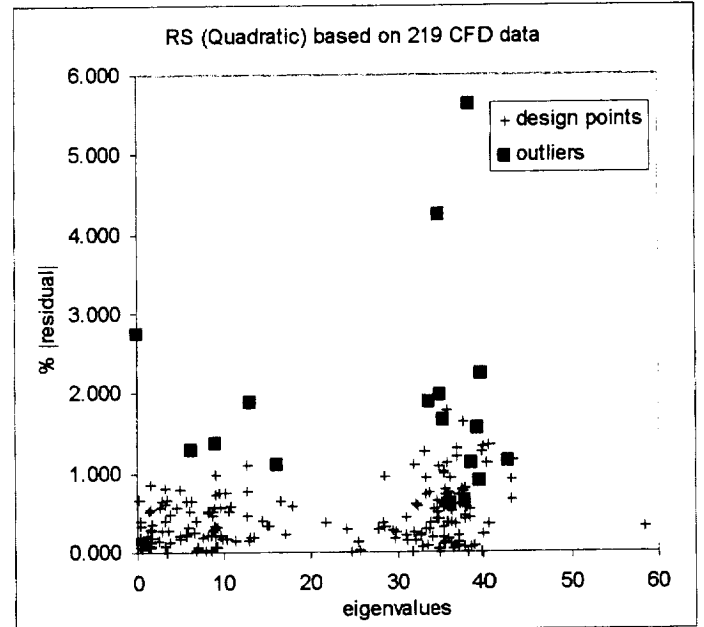
**Figure 1.** Leverage plots for Efficiency with 3 different models based on CFD-Data only for the first vane



(a) Efficiency vs. Eigenvalues



(b) Efficiency vs. % |Error|



(c) Eigenvalues vs. % |Error|

**Figure 2.** Quadratic RS bias error analysis against cubic RS by **Mean Squared Error** criterion based on CFD-data for the **first vane**